

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 22 日現在

機関番号：31302

研究種目：若手研究(B)

研究期間：2012～2014

課題番号：24700906

研究課題名(和文) 科学技術文書の論理性を推敲できる文章作成教育システムの構築

研究課題名(英文) Development of an Education System for Polishing Logicality of Scientific and Technical Documentation

研究代表者

松本 章代 (MATSUMOTO, Akiyo)

東北学院大学・教養学部・講師

研究者番号：40413752

交付決定額(研究期間全体)：(直接経費) 1,600,000円

研究成果の概要(和文)：研究代表者は、数年前から「学生自身による推敲・校正を支援する教育システム」を構築している。本研究の目的は、このシステムに「論理性の観点から推敲を支援するための機能」を追加することである。まずは「文書全体の流れの可視化」機能を実現した。続いて、レポート内において論理的に問題のある個所を自動検出し、指摘する(再検討を勧める)機能を実現するために、実際の学生レポートと教員によるレポートとを比較しながら、問題が生じやすい表現を分析した。

研究成果の概要(英文)：A technical writing support system developed during the last several years can detect problems in students' papers automatically. We intend to expand its functionality. This report emphasizes analysis of 'connective expressions' and 'sentence final expressions', and analysis of their respective frequencies in papers. Those analyses evaluate differences between students' papers and teachers' samples.

研究分野：自然言語処理

キーワード：自然言語処理 文書作成法 テクニカルライティング 接続詞 論理的 テキストマイニング 文章教育 作文指導

1. 研究開始当初の背景

文章作成指導のもっとも有効な手段は、担当教員によるきめ細かい添削指導であると考えられる。しかしながら、大人数を対象とした授業において添削指導を行うとなると、教員の労力は膨大なものとなる。しかも、添削指導の際には、一度読み直せば気が付きそうな不注意によるミスや、ルールを知ってさえいれば防げる誤りが多く目につく。添削する立場としては、せめてこのようなミスは無い状態で提出してほしいと願うが、自分自身が書いた文章を客観的に見直すことは難しく不備・不具合に気が付きにくいものである。ルールを守った文章が書けるようになるのにも、完全に身に付くまである程度トレーニングが必要である。また、学生の多くは他人に読んでもらうための文章を書くという意識が希薄である、ということも問題の一因である。

そこで「自分自身が書いた文章を客観的に見直す手助けをする」「科学的文章のルールを身に付けさせる」ための機能を有し、「推敲・校正の必要性を理解させ、他者に提出する文章は十分に吟味する癖をつけること」を目的とした、「学生自身による推敲・校正を支援する教育システム」を数年前から構築している。さらに、研究代表者が所属する東北学院大学教養学部情報科学科で開講されている、1年生を対象とした必修科目「初年次教育」において、本システムを運用している。このシステムは、以下の機能から構成される。

- ・ 校正を支援する「科学的文章のルールチェック機能」
- ・ 簡潔性・一義性の観点から推敲を支援する「わかりにくい文の指摘・可視化機能」

研究代表者はこの文章作成法の授業において、簡潔性・一義性の他に論理性についても重要視した指導を行っているが、文書全体の論理性について推敲を行うための機能が存在しない。そこで、文書全体の論理性をチェックできる機能を追加し、より効果的な教育システムとして完成させることを目指す。

2. 研究の目的

本研究の目的は、「学生自身による推敲・校正を支援する教育システム」に論理性の観点から「推敲」を支援するための機能を追加することである。

まずは「文書全体の流れの可視化」を実現する。

続いて、レポート内において論理的に問題のある箇所を自動検出し、指摘する(再検討を勧める)機能を実現する。

そのために、実際の学生レポートと教員によるレポートとを比較しながら、問題が生じやすい箇所を分析する。

3. 研究の方法

(1) 調査対象の文書

コーパスとして、2011年度・2012年度の

「初年次教育」において学部1年生が書いたレポートのうちワープロソフトで印刷されたものを用いる。レポートのテーマは2011年度・2012年度ともに(こちらが指定した、とある)ツールの機能の調査と評価、である。ただし、学生レポートの出来は「初年次教育」の授業内容にも影響されると考えられ、年度が異なると同一条件とは言えなくなることから、2011年度と2012年度を分けて分析する。以下、2011年度の70本を(A)、2012年度の84本を(B)とする。

一方、学生の書いたレポートの特徴をとらえるため、比較対象として大学教員が書いたレポート50本を用いる。これは、とある大学教育実践に関する研究発表会の予稿集の原稿であり、執筆者が大学教員に限定されているものである。1本あたりの分量はA4用紙2段組み2枚であり、学生が書くレポートの量と大きな差はない。以下、これを(C)とする。

なお、(A)(B)(C)いずれにも執筆者の重複はない。

(2) 調査の前提

著者は以前から、文書の内容・題材に依存しないながらも、その特徴を表すものとして、文頭の接続表現や文末表現に着目している。

接続詞は前後の文脈の関係を示す。読みやすい文章の全体構造を支えるのは、接続詞である。論理的な文書を作成するためには、適切な語を選択する必要がある。

一方、日本語は思考、状態、意志、態度など、文末に重要な表現が多くあらわれる。日本語の文の場合、最も文末にある自立語が文の中心要素であるから、その語と結びついて助動詞や助詞も含めて、この部分は文全体の意味に深くかかわることになる。すなわち、文末は適切な表現である必要がある。

このように、文頭の接続表現と文末表現はともに文書において重要な役割を担っており、学生のレポートにふさわしくない表現が存在している場合に、それを指摘することは有用である。教員によるレポートと学生のレポートとを比較すれば、レポートにふさわしい表現・ふさわしくない表現が特徴として現れることが期待できる。

なお、接続詞ではなく「接続表現」としたのは、文を接続する働きをするのは接続詞だけではないからである。文頭の表現によっては、接続詞ではなくても文と文をつなぐ機能を持っている場合がある。そこで本稿における「接続表現」は、以下の条件のどちらかを満たすものと定義する。

- ・ 文頭の接続詞
- ・ 文頭の副詞のうち文をつなぐ役割をもつもの

なお、文頭に限定する理由は、今回の調査が「接続詞の直前にある文で述べられている内容を受ける接続詞」を調査対象としたいためである。ただし、「文をつなぐ役割」をもつか否かについては、係り受け解析器のみで

は判定できないため最終的に研究代表者が判断を行う。

また、文末表現の定義は、「文の終端からさかのぼって自立語が出現するまで（自立語を含まない）」とする。

(3) 調査目的

教員・学生それぞれのレポートに含まれる文頭の接続表現および文末表現から、学生レポートの特徴を明らかにする。

また、学生の書いたレポートには、

- ・ 「よって」「したがって」といった因果関係を表す接続表現
- ・ 「つまり」「すなわち」などの言い換えを行う接続表現
- ・ 「このように」「こうして」といった結論を導く接続表現

を強引に使用し、論理の飛躍が起こっているケースが散見される。そこで、これらに該当する接続表現をまとめて本稿では「論理を導く接続表現」と呼び、さらにこれらに焦点を絞って分析を行う。これらの接続表現が論理的に不適切に用いられているケースに（適切な場合と区別できるような）何らかの共通点を探る。

(4) 調査手法

語句の頻度に基づく文書間の特徴分析

文書内に含まれる語句によって文書間の特徴を分析するための代表的な手法として、主成分分析と対応分析がある。文書間の特徴分析における主成分分析・対応分析では、語句の散布図と文書の散布図を重ね1つにしたグラフ（これをバイプロットという）に表現することにより、データの分類やデータ間の関係の把握を行うことができる。バイプロットによって情報が集約され、効率的な検証・考察が可能になる。

本研究においても、学生レポートと教員レポート間の特徴分析を行いたいため、主成分分析と対応分析を行う。

具体的な手続きについて、以下に示す。

- [1] 対象とするすべての文書に対し、株式会社 CSK によって開発された日本語係り受け解析器を用いて単語分解と品詞特定を行い、該当する表現（文頭の接続表現・文末表現）の抽出を行う。
- [2] (A)(B)(C)のグループごとに、各表現の出現頻度を算出する。
- [3] 統計解析ツール R を用いて、出現頻度に対し主成分分析および対応分析を行う。
- [4] それぞれバイプロットを確認し、図から学生レポートの特徴を読み取る。

「論理を導く接続表現」前後の論理性

まず、(A)(B)に含まれるすべての「論理を導く接続表現」について、第3者によってその出現付近の文章を読み、論理的に問題のあるケースとないケースに分類してもらう。

つぎに、非論理的と認められるケースについて、接続表現ごとに集計し、使用頻度の高いものを対象として、論理性の自動判定に利用できそうなルールを検討する。

4. 研究成果

(1) 文書全体の可視化

論理性が求められる技術文書では、筋がとおる順序で述べていくことが重要である。順序が不自然であったり接続詞が省略されたりすると論旨が理解しにくくなる。文の欠落（論理の飛躍）や接続詞の不適切な使用などがあれば、論理性はさらに損なわれる。

そこで、章または節単位の文集合全体の可視化を行う。「余計な修飾表現は無い方が、話の流れに集中しやすく、文と文の関係を見直す作業の支援につながる」という仮説に基づき、文書全体のあらすじを可視化することにより、論理展開のチェックを支援する。文の骨組みと論理チェックに必要な接続詞・接続助詞を残し、余計な修飾語・句や連体修飾節をそぎ落とすことにより、文書全体の筋を読み取ることが容易になると考える。抽出された語とその関係から図を生成する。学習者は表示された図を見て、筋がとおっているかどうか、確認を行う。特に、接続詞・接続助詞に着目させる。論理の飛躍はないか、文の順序は適切か、接続詞の不足はないか、などに注意しながらチェックを行う仕組みである。

(2) 非論理的個所の検出

語句の頻度に基づく文書間の特徴分析の結果

まず、(A)(B)(C)それぞれの文書数、総文数、総語数(a)、を表1に示す。

以降、各表現の出現回数(b)は10万語あたりの調整頻度($b/a*100,000$)で扱う。

表1. 調査対象文書の基本情報

	(A)	(B)	(C)
文書数	70	84	50
総文数	2,797	3,979	3,149
総語数(a)	44,212	67,394	76,611

➤ 論理を導く接続表現の分析

接続表現全体の出現頻度の合計は、(A)が526、(B)が1065、(C)が649である。それに対し、論理を導く接続表現のみの場合は、(A)が90、(B)が183、(C)が63である。論理を導く接続表現は特に学生レポートに多く含まれる傾向にあることがわかる。説明が不十分な状態で結論を急ぎ、論理の飛躍が起こっている可能性があると考えられる。

続いて、(A)(B)(C)を変数、論理を導く各接続表現の頻度をケースとして、主成分分析を行う。その結果、固有値は第1主成分が2.20、第2主成分が0.63、寄与率は第1主成分が71.57%、第2主成分が22.35%となった。第2主成分と合わせると累積寄与率は93.92%になることから、この2つの成分で元のデータの多くが説明できているといえる。

第1主成分と第2主成分のバイプロットを図1に示す。まず、変数に着目すると、縦軸において(C)は負の方向、(A)(B)は正の方向を向いており、しかも近接している。このこ

とから、(A)(B)にはかなりの類似性が認められる一方、(C)とは明らかに使用される接続表現に差があることがわかる。次に、ケースに着目すると、横軸の値が小さいものほど出現頻度が高く、「つまり」や「このように」「よって」などが特に学生に好んで使われる傾向にあることがわかる。

さらに、対応分析を行い、その結果のバイプロット(図2)に示す。図中の(A)(B)の間に、学生レポートに特に出現しやすい接続表現が表れている。

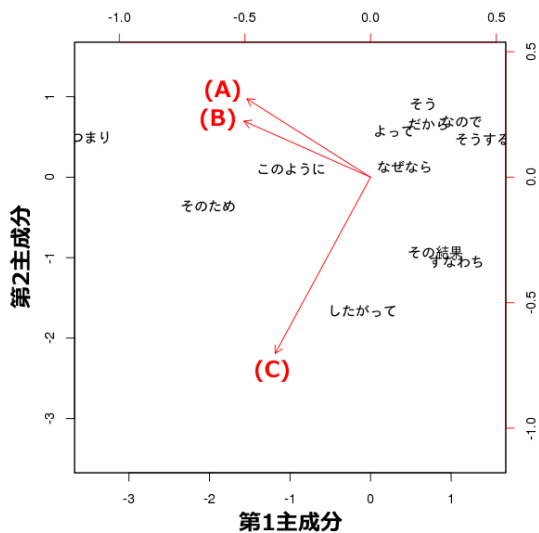


図1. 論理を導く接続表現の主成分分析のバイプロット

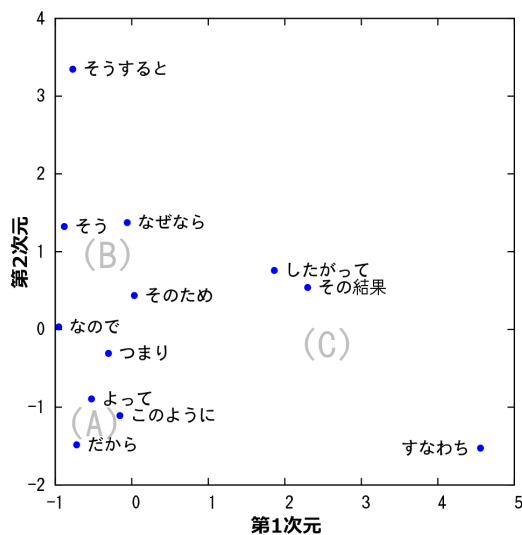


図2. 論理を導く接続表現の対応分析のバイプロット

「論理を導く接続表現」前後の論理性的の調査結果

(A)(B)に含まれる各「論理を導く接続表現」についてその前後の文脈の論理性を確認したところ、少なくとも(A)は38件、(B)は48件が非論理的であると認められた。接続表現別に集計すると、86件中、1位が「つまり」

で16件、2位が「そのため」で15件を占めた。そこで、「つまり」「そのため」に着目し、非論理的な使われ方をしているケースに共通し、適切な使われ方をしているケースには該当しないような規則性を探した。

その結果、「つまり」については有用な知見を得られなかったが、「そのため」については次の知見を得られた。非論理的であると判断される「そのため」を含む文の直前には、根拠とは関係ない文や意見・推測の文が存在している、という特徴が見受けられた。そこで、「そのため」を含む文の2文前までに「～だろう」や「～と思う」といった意見・推測の文が存在しているケースをプログラムによって自動抽出したところ、それに該当するものが学生レポートから8件抽出された。この8件はすべて先述の調査において、根拠が不十分で論理が飛躍している、あるいは意味が曖昧という理由で、「非論理的」と判断されたものであった。

考察

論理を導く接続表現の分析からは、学生は論理を導く接続表現の使用頻度が高いこと、学生が好んで用いる表現と一般的なレポートに多く用いられる表現は、やはり異なる傾向にあることがわかった。

「論理を導く接続表現」前後の論理性的の調査からは、非論理的な展開に用いられている接続表現は1位「つまり」、2位「そのため」という結果が得られた。特に「そのため」については、直前に意見・推測の文があると非論理的になる、という傾向が認められた。3位以下は件数が少ないため十分な検証ができないが、同様の特徴をもつ接続表現は他にもあると考えられる。一方、「そのため」が非論理的になりやすい理由として、「そのため」に2つの意味があることも影響していると考えられる。ひとつは、接続詞「したがって」「ゆえに」「よって」などとほぼ同じ意味で、因果関係を表す。もうひとつは、「～をするために」という意味である。目標や目的を達成するための方法や手段を表現する。それゆえ「そのため」は曖昧になり易い。もちろん文脈によって指示語が何を示すかを推測できることも多いが、原則としてレポートの文章は一義的であることが望ましい。

(3) まとめ

「学生自身による推敲・校正を支援する教育システム」に、「論理性的の観点から推敲を支援するための機能」を追加するため、まずは「文書全体の流れの可視化」機能を実現した。続いて、レポート内において論理的に問題のある箇所を自動検出し、再検討を勧める機能を実現するために、実際の学生レポートと教員によるレポートとを比較しながら、問題が生じやすい表現を分析した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 3件)

1. 松本章代: 他文書との比較に基づく学生レポートにおける問題点の解析, 人間情報学研究, Vol. 20, pp. 29-38 (2015.03). 査読有
2. 高橋光一, 佐藤 篤, 松本章代, 牧野悌也, 星野真樹: 情報科学科における文章授業 - 報告と課題 -, 東北学院大学教養学部論集, No. 169, pp. 29-47 (2014.12). 査読無
http://www.tohoku-gakuin.ac.jp/research/journal/bk2014/pdf/no09_03.pdf
3. 松本章代: 科学的文章の推敲・校正を支援する教育システムの構築, 東北学院大学教養学部論集, No. 167, pp. 53-62 (2014.03). 査読無
http://www.tohoku-gakuin.ac.jp/research/journal/bk2014/pdf/no01_05.pdf

〔学会発表〕(計 17件)

1. 柳沢雪絵, 松本章代, 佐伯 啓: Skype通話を利用した外国語会話訓練システムの構築と運用, 電子情報通信学会 2015年総合大会 ISS 特別企画 学生ポスターセッション (2015.3.11) 立命館大学.
2. 柳沢雪絵, 松本章代, 佐伯 啓: Skype通話を利用した外国語会話訓練システムの改善と運用, 平成 26 年度 情報処理学会東北支部研究会 (2015.02.10) 東北学院大学.
3. 松本章代, 斎藤 亮汰, 佐伯 啓: 大学生の文章表現力向上を目指した授業を支援する投稿・批評システムの開発と運用, 情報処理学会 第 14 回教育学習支援情報システム研究発表会 (2014.10.25) 東北大学.
4. 松本章代, 菅原 研, 武田 敦志, 松尾 行雄, 星野 真樹, 坂本 泰伸, 乙藤 岳志, 土橋 宏康, 松澤 茂: 予習・復習に重点をおいた初級プログラミングの授業実践 より有効な反転授業を目指す, 平成 26 年度 教育改革 ICT 戦略大会 (2014.09.05) 私学会館.
5. 大友麻実, 松本章代: 学生レポートにおける接続表現に関する調査, 情報処理学会 第 76 回全国大会 (2014.03.13) 東京電機大学.
6. 木村 実穂, 松本章代, 佐伯 啓: Skype通話を利用した外国語会話訓練システムの開発と教育効果の検証, 平成 25 年度 情報処理学会東北支部研究会 (2014.02.13) 東北学院大学.
7. 梁川 裕太, 松本章代: Android タブレット端末向け添削支援アプリケーションの開発, 平成 25 年度 情報処理学会東北支部研究会 (2014.02.13) 東北学院大学.
8. 大友 麻実, 松本章代: 学生のレポートにおける非論理的接続表現の検出, 平成

25 年度 情報処理学会東北支部研究会 (2014.02.13) 東北学院大学.

9. 松本章代, 大友麻実: 学生のレポートにおける非論理的接続表現の分析, 情報処理学会自然言語処理研究会第 215 回研究発表会 (2014.02.06) 国立情報学研究所.
10. 松本章代, 木村実穂, 佐伯 啓: 外国語会話訓練を目的とした Skype 音声配信システムの開発, 教育システム情報学会 (JSiSE) 2013 年度第 3 回研究会 (2013.09.14) 静岡大学
11. 松本章代, 菅原 研, 武田 敦志, 松尾 行雄, 星野 真樹, 坂本 泰伸, 乙藤 岳志, 土橋 宏康, 松澤 茂: 予習・復習に重点をおいた初級プログラミングの授業実践, 平成 25 年度 教育改革 ICT 戦略大会 (2013.09.05) 私学会館.
12. 松本章代, 木村 実穂, 佐伯 啓: Skype を利用した外国語会話訓練システムの構築, 情報処理学会コンピュータと教育研究会第 120 回研究発表会 (2013.07.06) 青山学院大学.
13. 松本章代, 今村 真浩, 小西 達裕, 高木 朗, 小山 照夫, 三宅 芳雄, 伊東 幸宏: 日本語ウェブ文書を対象とした 10 年間の実態調査 (2002 年 ~ 2012 年), 第 5 回データ工学と情報マネジメントに関するフォーラム (2013.03.04) 磐梯熱海ホテル華の湯.
14. 藤島 洋斗, 松本章代: C 言語プログラムを対象とした剽窃検出システムの開発, 平成 24 年度 情報処理学会東北支部研究会 (2013.02.13) 東北学院大学.
15. 浅井 彬弘, 松本章代: タブレット端末向け添削アプリケーションの開発, 平成 24 年度 情報処理学会東北支部研究会 (2013.02.13) 東北学院大学.
16. 青木大輔, 松本章代, 高橋光一: 学生レポートにおいて不適切な使われ方をしている接続詞の検出, 教育システム情報学会 2012 年度第 5 回研究会 (2013.01.12) 東北大学.
17. 松本章代, 高橋光一: 科学技術文書の論理性を推敲できる文章作成教育システムの構築, 教育システム情報学会 第 37 回全国大会 (2012.08.23) 千葉工業大学.

〔その他〕

ホームページ等

技術文書校正・推敲支援システム

<http://mmt1.cs.tohoku-gakuin.ac.jp/twss/>

6. 研究組織

(1) 研究代表者

松本章代 (MATSUMOTO Akiyo)
東北学院大学・教養学部・講師
研究者番号: 40413752

(2)研究分担者
なし

(3)連携研究者
なし

(4)研究協力者
高橋 光一 (TAKAHASHI Koichi)
東北学院大学・教養学部・非常勤講師
研究者番号：20128961